Mental Phenomena as Causal Determinants in Brain Functions

R. W. Sperry

Editor's Note: Although R. W. Sperry does not write from the perspective of process philosophy, his research on brain functions has led him to an emergent, interactionist view of mental phenomena which is highly congenial to Whitehead and Hartshorne. In April, 1974, and February, 1975, he has discussed these views in meetings sponsored by the Center for Process Studies in Claremont. The rich experimental evidence he amasses in support of his theory will prove of particular interest to process philosophers interested in the traditional mind/body problem. The following article is an abridged version of a paper read at a conference for Dialogues on Brain and Consciousness in April, 1973, at the University of California at Irvine. The author's work is supported by USPHS Grant No. MH 03372 and the Hixon Fund of the California Institute of Technology.

I. Introduction

The central concepts concerning consciousness that I shall try to defend in my part of this dialogue have already been presented in some detail (15; 16; 17; 18; 20; 21; 22). I shall accordingly review them only in brief outline, devoting the bulk of the discussion to various peripheral aspects and implications that have previously had less emphasis. At the outset let me make it clear that when I refer to consciousness what I mean is that kind of experience that is lost when one faints or sinks into a coma. It is the subjective experience that is lacking during dreamless sleep, that may be obliterated by a blow on the head, by anoxia, or by pressure on the inner walls of the third ventricle during brain surgery. On the positive side we can include as conscious events the various sensations elicitable by a local electric current applied to the unanesthetized brain, or the pain of a phantom amputated limb, as well as most of our waking subjective experience, including self-consciousness.

I want to emphasize, however, that I shall not be concerned particularly with *self*-consciousness any more than with the consciousness of other selves or with that of external objects, situations, and events. *Self*-consciousness is a separate story in itself. Neither shall I be trying to define different forms of consciousness, nor intermediate states between

full awareness and the *sub*conscious or the *un*conscious. My arguments can all be referred to some clearly accepted and preferably simple example of conscious experience: like seeing red, or hearing a musical tone, or feeling pain. The problem is difficult enough in its simplest and clearest formulation without introducing the confusion of borderline states. I assume that if we can find an answer to the mind/brain problem in its simplest form, we shall then be able to apply the basic concepts to the more complex aspects.

For the sake of further clarification let me specify that I shall address myself throughout to the problem of the nature of consciousness and the mind/brain relation as it is present *in other people's brains* primarily, rather than in my own brain. This hopefully avoids various logical entanglements that arise otherwise. This starting move is based on the assumption, of course, that other people's brains do have consciousness much like my own. Those who are not willing to accept this assumption have, I suspect, a separate problem all their own. I am not trying to avoid entirely by this step questions concerning the privacy of conscious experience. A number of different approaches to this important privacy, or first-person, property of consciousness are recognized, and I will try to outline later in context the explanation to which my own position leads.

Perhaps the quickest way to center in on our current interpretation is to compare it broadly with others. We can start by saying that our does not belong among positions based on dualism, epiphenomenalism, or other parallelisms. We can bypass as well the radical behaviorist refusal to consider the problem and also various sophistries and epistemological gymnastics that would make it just a pseudoproblem or explain it away as unimportant or nonexistent. We can also bypass the traditional materialism of the hard-core reductionistic and dialectic varieties. Our position does not accord either with the interpretation of subjective experience as just an inner aspect of the one material brain process. It is further distinguishable from the so-called "identity theory," that version of materialism which stresses that mental pehenomena are *identical* with the neural events (3). This view does not correlate consciousness with language particularly. Finally, it is in disagreement with the position known as panpsychism in which every individual entity in the universe is held to possess consciousness or psychic properties of some sort.

II. AN EMERGENT THEORY

On the positive side our present view can be classified broadly as an "emergent" theory of mind that needs to be distinguished from other emergent theories which have been advanced previously mainly by the Gestalt school in psychology. It differs from these latter in several respects. First, the phenomena of subjective experience are not thought to be derived from electrical field forces or volume conduction effects, or any metaneuronal by-product of cerebral activity. Our view relies on orthodox neural circuit and related physiological properties (9; 17). Second, there is no assumption of an isomorphic or topological correspondence between the events of perceptual experience and corresponding events in the brain.

I have conceived the mental properties to be *functional* derivatives that get their meaning from the way in which the brain circuits and related processes operate and interact rather than in terms of isomorphic correlations (17; 19). Reference to "spatiotemporal patterning" of brain activity is safe as far as it goes, but this term fails to connote the operational derivation of the conscious properties that I have tried to emphasize. Third, the conscious subjective properties in our present view are interpreted to have causal potency in regulating the course of brain events. The mental forces or properties, that is, exert a regulative control influence in brain physiology. The subjective conscious experience on these terms becomes an integral part of the brain process rather than a correlated phenomenon as conceived by Köhler (4) and others. The mental events are *causes* rather than *correlates.* In this respect our view can be said to involve a form of mental interactionism except that there is no implication of dualism or other parallelism in the traditional sense. The mental forces are direct causal emergents of the brain process.

When I initially stated this view in 1965, one had to search a long way in philosophy and especially in science to find anyone who would put into writing that mental forces or events are capable of causing physical changes in an organism's behavior or its neurophysiology. With rare exceptions, writings in behavioral science dealing with perception, imagery, emotion, cognition, and various other mental phenomena were very cautiously phrased to conform with prevailing materialist-behaviorist doctrine. Care was taken, that is, to be sure that the subjective phenomena should not be implied to be more than passive correlates or inner aspects of brain events and especially to avoid any implication that the mental phenomena might interact causally with the physical brain process. Those few who did earlier subscribe to psychophysical interaction were such extreme dualists that little heed was paid to them in behavioral science. Once we could show how mental events can causally influence neural events in a compromise formulation that does not violate the principles of scientific explanation, the long-standing resistance to mental/physical interaction began to decline. It is only since then that mental imagery, for example, has been able to gain popular acceptance as an explanatory construct. Today it becomes increasingly difficult to differentiate some of the closely related positions on these matters, and one must go back to the "pre-'65" versions in order to make clear distinctions.

The stand is taken here that wholes and their properties are real phenomena and that these and their causal potency are just as important as the properties of the parts to which the reductionist position likes to give prior, or even sole recognition. This is to say that the relationships of the parts to each other in time and space are of critical importance in causation and in determining the nature and properties of all entities. It is a pragmatic interpretation of what is real and meaningful. The properties of the parts are themselves in turn, of course, holistic properties of subsystems at a different level. The reductionist approach that would always explain the whole in terms of the parts leads to an infinite nihilistic regress in which eventually everything is held to be explainable in terms of essentially nothing.

III. A SIMPLE APPROACH

The way in which mental phenomena are conceived to control the brain's physiology can be understood very simply in terms of the chain of command in the brain's hierarchy of causal controls (15). It is easy to see that the forces operating at subatomic and subnuclear levels within brain cells are molecule-bound and are superseded by the encompassing configurational properties of the brain molecules in which the subatomic elements are embedded. The nuclear and other subatomic elements, that is, are pushed and hauled about in chemical interactions by the enveloping molecular properties. In the same way the properties of the brain molecules are enveloped by the dynamics of cellular organization, and the properties of the brain cells are in turn superseded by the larger network properties of the circuit systems in which they are embedded.

At the apex of the brain's organizational hierarchy are found the large cerebral processes that mediate mental activity. These large cerebral events as entities have their own dynamics and associated properties that causally determine their interactions. These top level systems' properties supersede those of the various subsystems they embody.

Only *some* of the dynamic holistic properties that emerge in the higher levels of cerebral activity are conscious phenomena. Many others, however, are not, even though the unconscious activities may in some cases be equally or more complex. Complexity alone is not, in our scheme, the source of the conscious qualities (9). It is the operational function rather than the complexity of any given cerebral process that determines its conscious effect.

A mutual interdependence exists between the neural events and the emergent mental phenomena. In other words the brain psysiology determines the mental effects, and the mental phenomena in turn have causal influence on the neurophysiology. The interjection of subjective mental experience into the causal sequence of decision-making on these terms brings a compromise, not only between materialism and mentalism, but also between the positions of determinism and free will. Determinism of this kind, in which subjective experience is included as a causal agent in brain function, allows degrees of freedom in any voluntary choice far above that envisaged in traditional materialist or atomistic determinism.

I have tried to tie these general principles to the example of subjective pain as it is referred to an amputated limb (15). For present purposes let us make it more specifically the pain of a phantom left foot that is produced by stimulation of a sore toe in the opposite hindfoot in one of our experimental "sensory nerve cross" rats. These are rats in which the right hindfoot has become reinnervated by foreign nerves that originally had supplied the left foot (11). The switch in nerve connections from left to right foot is brought about by surgical cross-union of the sciatic nerve and its branches from left to right leg in the fourth week after birth as a test of central nervous plasticity and the functional interchangeability of nerve connections. Occasionally the animals will "instinctively" chew off the denervated insensitive foot on the left, and there is also a tendency for cutaneous trophic sores to develop in the right foot while it is being

reinnervated. Such a sore on the right foot heals very slowly despite antibiotics because these rats walk around on three legs, protectively holding up the wrong foot from which the pain seems to come and thereby putting additional pressure and trauma on the sore right foot. Occasionally, as the result of an extra hard impact or abrasion to the right foot, the rat may yip or squeak and will turn to lick, not at the sore right foot, but at the uninjured left foot when it is there, or otherwise at the amputation stump.

I choose this example to emphasize among other things my assumption that conscious experience is not restricted to the human species. Self-consciousness is another matter, of course, and may well be limited mainly to man with some beginnings in the higher subhuman forms. The experimental rat's false reference of pain to the amputated left foot persists throughout life, and this example thus serves also to reinforce the inference that the basic circuit properties responsible for conscious experience are genetically determined in large part. They may have evolved initially around sensory functions or/and around a primitive awareness with positive and negative reinforcement functions.

The main point to be brought out with this example is the contention that the animal's responses in protectively holding up the wrong foot and in yipping and licking the wrong foot are caused directly in brain function by the subjective pain property itself rather than by the physiology of the nerve impulses, or by the chemical, atomistic, or other subunit features of the brain process. The pain sensation is considered to be a real emergent phenomenon in itself. Although built of neural events and possibly of glial events as well, the pain sensation as a larger whole is not itself the same as the constituent neural and glial events. Nor is the subjective pain to be viewed either as a mere parallel correlate of the brain process. Rather I look upon it as a real dynamic entity in the brain activity that has an important causal role as a phenomenon itself in the stimulus-response sequence. In other words, a full objective account of the whole stimulus-response process would not be complete without including the pain as such. Although our neurophysiology is not yet sufficiently advanced to give an adequate description of the neural composition of the pain phenomenon, or of other conscious events, one assumes that this probably will be possible eventually as our knowledge of brain mechanisms continues to advance.

IV. The Bisected Brain and Unity of Consciousness

Philosophy has been concerned with the "unity of consciousness" in connection with problems relating to the nature of the self, the person, and personal identity. In our "split-brain" studies of the past two decades (9; 10; 12; 13; 14; 20) the surgically separated hemispheres of animals and man have been shown to perceive, learn, and remember independently, each hemisphere evidently cut off from the conscious experience of the other. In man the language-dominant hemisphere further reports verbally that it is not consciously aware of the concomitant or immediately preceding mental performances of the disconnected partner hemi-

sphere. These test performances of which the speaking hemisphere remains unaware obviously involve perception, comprehension, and in some cases nonverbal memory, reasoning, and concept formation of different kinds, depending on the nature of the test task. In these and in many other respects, the split brain animal and man behave as if the separated hemispheres had each a mind of its own.

This division by surgery of the normally unified realm of conscious awareness into two distinct domains of conscious experience that exist in parallel and in some cases have content that is mutually contradictory has been subject to several different philosophical interpretations. One line of reasoning concludes that each hemisphere of the brain must have a mind of its own, not only after surgery but also in the normal intact state as well. The normal individual, that is, is interpreted to be a compound of two persons, one based in each hemisphere (1; 8). A contrasting interpretation says that only one, the language-dominant hemisphere, remains conscious (2:73-80), and thus the unity of consciousness is preserved. It is inferred that the disconnected minor hemisphere operates like an automaton or complex computer. Another view holds that consciousness is not centered in either right or left hemisphere but in some unified metaorganizing system (5), presumably in the intact brain stem. There are additional variations on these main themes (6).

My own inclination is to see consciousness as being unified in the normal brain but largely divided in the bisected brain, depending on the depth and extent of the surgery and depending also on the nature and level of the particular conscious process in question. I would credit the neocommissures with a unifying role in conscious activity under normal conditions that in effect serves to tie the conscious function of each hemisphere together across the midline into a single unified process. The callosal activity thus becomes part of the conscious event. The fiber systems uniting right and left hemispheres are viewed as being not essentially different in their relation to consciousness from those uniting front and back or other areas within the same hemisphere. This interpretation does not exclude the possibility that the conscious processes in left and right hemispheres may function separately in the undivided brain under exceptional conditions and particularly where pathology tends to depress commissural function.

Surgical separation of the hemispheres, especially the deeper bisections we perform in animals, I have interpreted as resulting in the creation of two distinct domains of consciousness. This says nothing about *self*-consciousness. It remains to be determined how much, if any, self-consciousness is present in the disconnected minor hemisphere of man. However, preliminary findings from experiments in progress in collaboration with Zaidel support the conclusion that the disconnected minor hemisphere does in fact exhibit characteristic self-conscious reactions to pictures of itself, showing appropriate emotional displays in different contexts.

This interpretation does not preclude retention in the bisected brain of a right-left unity in some aspects and levels of conscious experience. This is assured in part by bilateral sensory representation in each hemisphere as in the case, for example, with facial sensibility. We presume,

however, by extrapolation, that these unified "whole face" experiences in each hemisphere are cut off from their counterparts in the opposite hemisphere. The structure of the conscious cerebral process is inferred to be such that some aspects of conscious experience may be separated by commissurotomy, while others, united through bilateral representation and/or brain stem mechanisms, remain intact (13; 14; 15). In most of our work we have naturally emphasized the more interesting and striking aspects of consciousness that are separated by the surgery and which predominate in the kinds of test tasks we employ. However, I have also tried to stress the presence of many unifying factors. The possibility remains that some elemental components of consciousness stay unified in the split brain even in those tests where the bulk of the conscious content is clearly divided.

On these terms, neural activity transmitted through the corpus callosum becomes part of the conscious brain process. However, in order to properly comprehend the critical holistic properties of the conscious process one would have to include the associated activity on both sides. In the callosal fiber systems and these assorted cortical mechanisms on either side we probably come as close as anywhere in the brain to a direct grip on psychoneural relations. Consider, for example, the normally unified perception of the whole visual field and its division down the vertical midline that is produced by midline commissurotomy.

As knowledge of brain function and the mind/brain relation advances, one would anticipate that terms like "mind" and "person" will have to be redefined, or at least more precisely defined. Already it makes little sense, employing past definitions, to argue about how many "minds" or "persons" are present in the bisected brain. What is most needed is better understanding of the functional relationships between the neural mechanisms that are divided and those that are not, and their respective roles in the generation of conscious experience.

The neural mechanisms by which the mental effects in each hemisphere are generated may have common undivided brain stem and perhaps cerebellar components that may or may not have any conscious properties in themselves but which are essential substructure constituents of the conscious experience. Particularly important among the undivided brain stem components are the neural mechanisms of attention. Thus if one were to diagram schematically the structure of mind after cerebral commisurotomy, it would be crudely Y-shaped, containing a common stem with left and right upper arms in each hemisphere. Each hemisphere contains the prepresentation of a bilateral body schema in which the ipsilateral limb extremities are present but fainter and more crudely depicted. The external surroundings also have bilateral representation. This is much better for the contralateral side, especially in vision, but the ipsilateral half of space is not absent. Thus each disconnected hemisphere contains the anatomical substrate for a unified self in a bilateral surround. Each hemispheric representation is based in and functionally dependent upon intact brain stem mechanisms that are in part bilateral and which, of course, remain intact in the human commissurotomy patients.

One can ask, what separates the conscious part of the brain process from its lower level nonconscious foundations? Also, for any given stimulus-response sequence, what separates the nonconscious sensory input on the one side and the motor output on the other from the more central conscious portion of the total activity? Similarly, among the higher cerebral functions, what kind of boundary or interface do we picture between processes that have conscious properties and those that do not? The answer is that we do not picture anything separating the conscious from the unconscious neural events—aside from organizational coherence. No interface or other definite boundary is imagined to be interposed between the two.

Although the emergent properties are spoken of as encompassing or enveloping the constituent neural events, the implication is not that of an enveloping surface film or electrical potential difference or other interface, but only that of smaller neural events being caught up in the dynamics of larger neural events. A cerebral process acts as a conscious entity, not because it is spatially set apart from other cerebral activity, but because it functions organizationally as a unit. Presumably the conscious process may be interwoven with and may share active components with other brain processes that do not reach conscious awareness. The holistic properties are not to be conceived in simple spatial, volumetric, or dimensional terms but rather in terms of nerve network and cerebral circuit interactions, the emergent dynamics of which have yet to be elucidated, especially for the upper, conscious levels of brain function.

Normally, with the neocommissures intact, neural events in right and left upper arms of our schematized *Y* substrate of consciousness become merged into a unified conscious brain process. The criterion for unity is an operational one; that is, the right and left components, coalesced through commissural communication, function in brain dynamics as a unit. This is illustrated in the unified visual perception of a stimulus figure flashed tachistoscopically, half in the left and half in the right visual half fields. In the normal brain the right and left hemispheric components combine and function as a unit in the causal sequence of cerebral control. In the divided brain, on the other hand, each hemispheric component gets its own separate causal effect as a distinct entity.

V. IMPLICATIONS

Some of the main implications derive from the fact that conscious experience is given an operational causal role in cerebral function and hence a reason for being and for having been evolved. This is not true for prior materialistic or various parallelistic interpretations in which the brain would function just as well in terms of the neural events, whether or not these had subjective properties.

Although it is not difficult, as indicated (18), to stretch either the materialist or opposing mentalist approaches of ten years ago to incorporate these compromise psychophysical interaction concepts, it is important to recognize the various changes this involves. These changes have important consequences in other areas of philosophy that deal with de-

terminism and free will (15; 21), with the concept of causation (7), and with the whole field of human values and the relation of scientific explanation to value judgment (22). Value theory has been rather neglected in philosophy of late, but could take on new importance on our present terms especially in view of the critical significance of human value priorities in the context of mounting global crises.

Introduction of mental phenomena into the causal sequence of brain function means, among other things, that values of all kinds, even aesthetic, spiritual, and irrational, must now be recognized as positive causal factors per se in human decision-making. Our seemingly free choices remain causally determined, but at a mental rather than molecular or neuronal level of causation—literally by what we wish and most value. Considered broadly, the present interpretation goes far to restore to the scientific view of human nature some of the freedom of choice, dignity, creativity, personal responsibility, and other humanistic attributes of which it has long been deprived.

Our current interpretation leads to a unifying concept of mind, brain, and man in nature and points to a "this world" framework for human values—a framework within which science can function. Subjective values on our present terms become objective causal agents operating in the physical brain and through the brain onto the surrounding world. As the brain process comes to be understood objectively, all mental phenomena, including the generation of values, can be treated as objective causal agents in human decision-making. The origins, directive potency, and consequences of values all become amenable, in principle, to objective scientific investigation and analysis. This applies at all levels, from that of the pleasure-pain centers and other reinforcement systems of the brain, on up through the forces that mold priorities at the societal, national, and international plane. A science of values becomes theoretically feasible and a matter of top priority considering the critical role played by the human value factor in determining world crisis conditions.

REFERENCES

1. Bogen, J. E., "The Other Side of the Brain (II. An Appositional Mind)," *Bulletin of the Los Angeles Neurological Society* 34 (1969), 135-62.
2. Eccles, J. C. *Facing Reality: Philosophical Adventures by a Brain Scientist.* New York: Springer-Verlag, 1970.
3. Feigl, H. *The Mental and the Physical.* Minneapolis: University of Minnesota Press, 1967.
4. Köhler, W., and R. Held, "The Cortical Correlate of Pattern Vision," *Science* 110 (1949), 414-19.
5. MacKay, D. M., "Discussion," *Brain and Conscious Experience,* ed. J. C. Eccles. Heidelberg: Springer-Verlag, 1966, pp. 312f, 422-44.
6. Nagel, T., "Brain Bisection and the Unity of Consciousness," *Synthese* 22 (1971), 396-413.
7. Pols, Edward, "Power and Agency," *International Philosophical Quarterly* 11 (1971), 293-313.

8. Puccetti, R., "Brain bisection and personal identity," *British Journal for the Philosophy of Science* (March, 1974).
9. Sperry, R. W., "Brain Bisection and Mechanisms of Consciousness," *Brain and Conscious Experience*, ed. J. C. Eccles. Heidelberg: Springer-Verlag, 1966, pp. 298-313.
10. Sperry, R. W., "Cerebral Organization and Behavior," *Science* 133 (1961), 1749-57.
11. Sperry, R. W., "Functional Results of Crossing Sensory Nerves in the Rat," *Journal of Comparative Neurology* 78 (1943), 59-90.
12. Sperry, R. W., M. S. Gazzaniga and J. E. Bogen, "Interhemispheric Relationships: The Neocortical Commissures," *Handbook of Clinical Neurology*, ed. P. J. Vinken and G. W. Bruyn. New York: John Wiley and Sons, 1969, vol. 4, pp. 273-90.
13. Sperry, R. W. *Lateral Specialization in the Surgically Separated Hemispheres:* The Neurosciences, Third Study Program, ed. F. O. Schmitt and F. G. Worden. Cambridge, Mass.: MIT Press, 1974.
14. Sperry, R. W. *Mental Unity Following Surgical Disconnection of the Cerebral Hemispheres.* Harvey Lectures. New York: Academic Press, 1968.
15. Sperry, R. W., "Mind, Brain, and Humanist Values," *New Views on the Nature of Man*, ed. J. R. Platt. Chicago: University of Chicago Press, 1965, pp. 71-92. Reprinted in the *Bulletin of the Atomic Scientists* (1966).
16. Sperry, R. W., "A Modified Concept of Consciousness," *Psychological Review* 76 (1969), 532-36.
17. Sperry, R. W., "Neurology and the Mind-Brain Problem," *American Scientist* 40 (1952), 291-312.
18. Sperry, R. W., "An Objective Approach to Subjective Experience: Further Explanation of a Hypothesis," *Psychological Review* 77 (1970), 585-90.
19. Sperry, R. W., and Nancy Miner, "Pattern Perception Following Insertion of Mica Plates into Visual Cortex," *Journal of Comparative and Physiological Psychology* 48 (1955), 463-69.
20. Sperry, R. W., "Perception in the Absence of the Neocortical Commissures," *Association for Research in Nervous and Mental Diseases* 48 (1970), 123-38.
21. Sperry, R. W., "Problems Outstanding in the Evolution of Brain Function." James Arthur Lecture. New York: American Museum of Natural History, 1964.
22. Sperry, R. W., "Science and the Problem of Values," *Perspectives in Biology and Medicine* (Fall, 1972). Reprinted in *Zygon* (1974).